# A SMOOTHING ALGORITHM FOR THE SMALLEST INTERSECTING BALL PROBLEM

Nguyen Thai An [1] and Nguyen Mau Nam[2]

**Abstract:** The Sylvester smallest enclosing circle problem asks for the smallest circle that encloses a finite number of points in the plane. Efficient algorithms for solving a generalized model of the classical Sylvester problem in which the given points are replaced by given Euclidean balls have been introduced and studied in the literature. In this paper, we present a fast numerical algorithm for solving another generalization of the classical Sylvester problem called the *smallest intersecting ball problem* which asks for a smallest Euclidean ball that intersects a finite number of convex sets.

**Key words.** MM Principle, Nesterov smooth algorithm, smallest enclosing circle problem, smallest intersecting ball problem

## 1 Introduction and Problem Formulation

The *smallest enclosing circle problem* asks for the circle of smallest radius enclosing a given set of finite points in the plane. This problem was introduced in the 19th century by Sylvester [10]. As one of the main problems of computational geometry, the problem has attracted great attention from many researchers. Current research focuses on developing fast numerical algorithms for solving the problem in large scale. In a recent paper, Nam et al. [6] proposed and studied the following generalized version of the smallest enclosing circle problem: given a finite number of nonempty closed convex subsets of a reflexive Banach space, find a ball with the smallest radius that intersects all of the sets. This new problem is called the *smallest intersecting ball problem*. Further generalized models of the smallest enclosing circle problem that involve both intersecting and enclosing generalized balls have been introduced and theoretically studied in [5].

In the general setting of [6], because of the intrinsic nondifferentiability of the data, the authors pursued the subgradient method to study the problem. It turns out that the subgradient method is applicable to a broad setting. However, the slow convergence rate prevents us from studying large-scale models. Our goal in this paper is to study the smallest intersecting ball problem in which the distance function is generated by the Euclidean norm of $\mathbb{R}^n$.

Let $\|\cdot\|$ denote the Euclidean norm in $\mathbb{R}^n$. Given a finite number of nonempty closed convex sets $\Omega_i$ for $i = 1, \ldots, m$, the mathematical modeling of the smallest intersecting ball

[1]Thua Thien Hue College of Education, 123 Nguyen Hue, Hue City, Vietnam (email: thaian2784@gmail.com)

[2]Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, Portland, OR 97202, United States (mau.nam.nguyen@pdx.edu). The research of Nguyen Mau Nam was partially supported by the Simons Foundation under grant #208785.

problem is

$$\text{minimize } \mathcal{D}(x) := \max\{d(x;\Omega_i) \mid i = 1,\ldots,m\} \text{ subjec to } x \in \mathbb{R}^n, \tag{1.1}$$

where the *Euclidean distance function* to a nonempty closed convex subset $Q \subseteq \mathbb{R}^n$ is given by

$$d(x;Q) := \inf\{\|x-q\| \mid q \in Q\}.$$

The *Euclidean projection* from $x$ to $Q$ is

$$\mathcal{P}(x;Q) := \{q \in Q \mid \|x-q\| = d(x;Q)\}.$$

It is well known that $\mathcal{P}(x;Q)$ is a singleton for every $x \in \mathbb{R}^n$.

Since all the sets $\Omega_i$ for $i = 1,\ldots,m$ are convex, the cost function $\mathcal{D}$ is convex but nonsmooth. Moreover, problem (1) can be equivalently formulated as follows

$$\text{minimize } \mathcal{D}_\gamma(x) := \max\{[d(x;\Omega_i)]^\gamma \mid i = 1,\ldots,m\} \text{ subjec to } x \in \mathbb{R}^n,$$

where $\gamma > 1$. In this new formulation, each function $\varphi_i(x) := [d(x;\Omega_i)]^\gamma$ is differentiable, but the new cost function $\mathcal{D}_\gamma$ is still nonsmooth. In the next sections, we only consider the case where $\gamma = 1$ although it is possible to develop an analog with $\gamma > 1$.

In this paper, we apply the smoothing technique developed in [12] along with the *MM Principle*, see for instance [2], and the *Nesterov smooth algorithm* [7] to provide a fast numerical scheme that is applicable for the smallest intersecting ball problem that involving a large number of sets in high dimensions.

If all of the target sets have a common point, then any point in common is a solution of problem (1), so we always assume that $\cap_{i=1}^n \Omega_i = \emptyset$. We also assume that at least one of the target sets $\Omega_i$ for $i = 1,\ldots,m$ is bounded, which guarantees the existence of an optimal solution. These are our standing assumptions in this paper.

## 2 Smooth Approximations

For $p > 0$, the *smoothing log-exponential function* of $\mathcal{D}$ is defined as follows

$$\mathcal{D}(x,p) = p \ln \sum_{i=1}^m \exp\left(\frac{g_i(x,p)}{p}\right), \tag{2.2}$$

where

$$g_i(x,p) = \sqrt{d(x;\Omega_i)^2 + p^2}.$$

**Lemma 2.1** *The function $\mathcal{D}(x,p)$ has the following properties:*
(i) *If $x \in \mathbb{R}^n$ and $0 < p_1 < p_2$, then*

$$D(x,p_1) < D(x,p_2).$$

(ii) *For any* $x \in \mathbb{R}^n$ *and* $p > 0$,

$$0 \leq \mathcal{D}(x, p) - \mathcal{D}(x) \leq p(1 + \ln m).$$

(ii) *For any* $p > 0$, $\mathcal{D}(\cdot, p)$ *is convex, coercive, and continuously differentiable with the gradient in* $x$ *computed by*

$$\nabla \mathcal{D}(x, p) = \sum_{i=1}^{m} \frac{\lambda_i(x, p)}{g_i(x, p)} (x - \widetilde{x}_i) \, .$$

*where* $\widetilde{x}_i := \mathcal{P}(x; \Omega_i)$ *and*

$$\tau(x, p) = \sum_{i=1}^{m} \exp\left(\frac{g_i(x, p)}{p}\right), \; \lambda_i(x, p) = \frac{\exp\left(\frac{g_i(x,p)}{p}\right)}{\tau(x, p)}.$$

**Proof.** The first and second assertion can be analogously proved as in [12]. For any $p > 0$, let $\varphi_i(x) = [d(x; \Omega_i)]^2$. Then $\nabla \varphi(x) = 2(x - \widetilde{x}_i)$, where $\widetilde{x}_i = \mathcal{P}(x; \Omega_i)$. Since every projection mapping $\mathcal{P}(\cdot; \Omega_i)$ is continuous, $\mathcal{D}(x, p)$ is continuously differentiable as a function of $x$. The function $f_p(u) = \frac{\sqrt{u^2+p^2}}{p}$ is increasing and convex on the interval $[0, \infty)$, and $d(\cdot; \Omega_i)$ is convex, so the function $k_i(x, p) = \frac{g_i(x,p)}{p}$ is also convex with respect to $x$. Take any $x, y \in \mathbb{R}^n$ and $\lambda \in (0, 1)$, by the convexity of the function $x = (x_1, \ldots, x_n) \to \ln(\sum_{i=1}^{m} \exp(x_i))$, one has

$$\begin{aligned}
\mathcal{D}(\lambda x + (1 - \lambda)y, p) &= \ln \sum_{i=1}^{m} \exp\left[k_i(\lambda x + (1 - \lambda)y, p)\right] \\
&\leq \ln \sum_{i=1}^{m} \exp\left[\lambda k_i(x, p) + (1 - \lambda)k_i(y, p)\right] \\
&\leq \lambda \ln \sum_{i=1}^{m} \exp\left(k_i(x, p)\right) + (1 - \lambda) \ln \sum_{i=1}^{m} \exp\left(k_i(x, p)\right) \\
&= \lambda \mathcal{D}(x, p) + (1 - \lambda)\mathcal{D}(y, p).
\end{aligned}$$

Thus, $\mathcal{D}(\cdot, p)$ is convex. Using the standing assumption stating that at least one of the target sets is bounded, without loss of generality, we assume that $\Omega_1$ is bounded. Then

$$\lim_{\|x\| \to \infty} \mathcal{D}(x, p) \geq \lim_{\|x\| \to \infty} d(x; \Omega_1) = \infty.$$

Therefore, $\mathcal{D}(\cdot, p)$ is coercive. The proof is now complete. $\square$

**Remark 2.2** (i) In general, $\mathcal{D}(\cdot, p)$ is not strictly convex. Indeed, in $\mathbb{R}^2$, we consider the problem with two target sets $\Omega_1 = \{-1\} \times [-1, 1]$ and $\Omega_2 = \{1\} \times [-1, 1]$. Then $\mathcal{D}(\cdot, p)$ takes constant value on $\{0\} \times [-1, 1]$.

(ii) If all of the target sets reduce to singletons, $\Omega_i = \{c_i\}$ for $i = 1, \ldots, m$, one has

$$\mathcal{D}(x, p) = p \ln \sum_{i=1}^{m} \exp\left(\frac{g_i(x, p)}{p}\right),$$

where $g_i(x,p) = \sqrt{\|x - c_i\|^2 + p^2}$, and the gradient of $\mathcal{D}(\cdot, p)$ at $x$ becomes

$$\nabla \mathcal{D}(x,p) = \sum_{i=1}^{m} \frac{\lambda_i(x,p)}{g_i(x,p)} (x - c_i) .$$

From [12] it follows that $\mathcal{D}(\cdot, p)$ is twice continuously differentiable and strictly convex. Moreover, for any $p > 0$, one has

$$\|\nabla \mathcal{D}(x,p) - \nabla \mathcal{D}(y,p)\| \leq \frac{2}{p}\|x - y\| \text{ for all } x, y \in \mathbb{R}^n.$$

# 3 The MM Principle

Consider the optimization problem

$$\text{minimize } f(x), x \in \mathbb{R}^n.$$

The main idea of the *MM Principle* is to find an appropriate approximation $\mathcal{M}$ of the objective function $f$ such that

$$\mathcal{M}(x,x) = f(x) \text{ and } f(x) \leq \mathcal{M}(x,y) \text{ for all } x, y \in \mathbb{R}^n.$$

Let $\mathcal{F}(x)$ the set of optimal solutions of the problem

$$\text{minimize } \mathcal{M}(x,y) \text{ subject to } y \in \mathbb{R}^n,$$

and call it the algorithm map. The MM algorithm is given by

$$x_{k+1} = \arg\min_x \mathcal{M}(x, x_k).$$

Finding an appropriate majorization is an important step in this algorithm. It has been shown in [2] that the MM Principle provides an effective tool for solving the *generalized Heron problem* introduced in [4]. The key step is to use the upper estimate

$$d(x;Q) \leq \|x - \mathcal{P}(y;Q)\| \text{ and } d(y;Q) = \|y - \mathcal{P}(y;Q)\|.$$

In what follows, we apply the MM Principle to solving the smallest intersecting ball problem in combination with the smoothing technique presented earlier.

# 4 Applications to the Smallest Intersecting Ball Problem

We would like to solve problem (1). In the first step, we approximate the cost function by the smoothing log-exponential function (2.2). Then the new function is majorized in order to apply the MM Principle. For $x, y \in \mathbb{R}^n$ and $p > 0$, define

$$G(x|y|p) := p \ln \sum_{i=1}^{m} \exp\left(\frac{\sqrt{\|x - \mathcal{P}(y;\Omega_i)\|^2 + p^2}}{p^2}\right).$$

Then $G(x|y|p)$ serves as a majorization of the smoothing log-exponential function (2.2), and hence of the cost function of problem (1).

Let us now present the *MM smoothing algorithm* for solving the smallest intersecting ball problem following the idea from [12].

**MM Smoothing Algorithm.**

INPUT: $m$ target sets $\Omega_i$, $i = 1, \ldots, m$.
INITIALIZE: $\sigma \in (0, 1)$, $y_0 \in \mathbb{R}^n$ and $p_0 > 0$.
Set $k = 0$.
**repeat**
Using a first order method to solve the problem
$\min_{x\in\mathbb{R}^n}\{G(x|y_k|p_k)\}$
and denote its solution by $x_k$.
Set $p_{k+1} = \sigma p$, $y_{k+1} = x_k$.
Set k:=k+1.
**until a stopping criterion is satisfied**.

To proceed, we present some basic concepts and results of convex analysis. A systematic development of convex analysis can be found, for instance, in [9]. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. For $\bar{x} \in \mathbb{R}^n$, a *subgradient* of $f$ at $\bar{x}$ is a vector $v \in \mathbb{R}^n$ that satisfies

$$\langle v, x - \bar{x}\rangle \leq f(x) - f(\bar{x}) \text{ for all } x \in \mathbb{R}^n.$$

The set of all subgradients of $f$ at $\bar{x}$ is called the *subdifferential* of $f$ at this point and is denoted by $\partial f(\bar{x})$.

From the definition, it is clear that $f$ has an absolute minimum at $\bar{x}$ if and only if $0 \in \partial f(\bar{x})$.

**Theorem 4.1** *Let $\{x_k\}_{k=1}^{\infty}$ be the sequence of exact solutions, $x_k = \arg\min_x G(x|y_k|p_k)$, generated as in the smoothing MM algorithm. Then any limit point of $\{x_k\}_{k=1}^{\infty}$ is an optimal solution of problem* (1).

**Proof.** Let $x_*$ be a limit point of the sequence $\{x_k\}_{k=1}^{\infty}$. Without loss of the generality, we assume that $x_k \to x_*$ as $k \to \infty$. By the definition, this also implies that $y_k \to x_*$ as $k \to \infty$. Because $x_k$ is a solution of the smooth minimization problem $\min_{x\in\mathbb{R}^n} G(x|y_k|p_k)$, one has

$$\nabla G(x_k|y_k|p_k) = \sum_{i=1}^{m} \frac{\lambda_i(x_k, p_k)}{g_i(x_k, p_k)} (x_k - \mathcal{P}(y_k; \Omega_i)) = 0,$$

where $g_i(x_k, p_k) = \sqrt{\|x_k - \mathcal{P}(y_k; \Omega_i)\| + (p_k)^2}$ and $\lambda_i(x_k, p_k) = \exp\left(g_i(x_k, p_k)/p_k\right)$. Denote $\lambda_i^k := \lambda_i(x_k, p_k)$ and let the active index set at $x_*$ be defined by

$$I(x_*) := \Big\{ i \in \{1, 2, \ldots, m\} \mid d(x_*; \Omega_i) = \mathcal{D}(x_*) \Big\}.$$

We can rewrite $\lambda_i^k$ as follows

$$\lambda_i^k = \frac{\exp\left(\frac{g_i(x_k,p_k)}{p_k}\right)}{\sum_{i=1}^m \exp\left(\frac{g_i(x_k,p_k)}{p_k}\right)} = \frac{\exp\left(\frac{g_i(x_k,p_k)-g_\infty(x_k,p_k)}{p_k}\right)}{\sum_{i=1}^m \exp\left(\frac{g_i(x_k,p_k)-g_\infty(x_k,p_k)}{p_k}\right)},$$

with $g_\infty(x_k,p_k) = \max\{g_i(x_k,p_k) \mid i = 1,\ldots,m\}$. By the continuity of the norm function and the Euclidean projection mapping to convex sets, we have

$$\lim_{k\to\infty} g_i(x_k,p_k) = d(x_*;\Omega_i), \text{and } \lim_{k\to\infty} g_\infty(x_k,p_k) = \mathcal{D}(x_*).$$

This implies

$$\lim_{k\to\infty} \lambda_i^k = 0 \quad \text{for } i \notin I(x_*). \tag{4.3}$$

By (4.3) and the fact that $\sum_{i=1}^m \lambda_i^k = 1$, $\lambda_i^k > 0$ for all $k \in \mathbb{N}$ and $i \in \{1,2,\ldots,m\}$, the bounded sequence $\{\lambda_i^k\}_{k=1}^\infty$, $i \in I(x_*)$, has a convergent subsequence. Without loss of generality, we suppose that

$$\lim_{k\to\infty} \lambda_i^k = \lambda_i^*, \quad i \in I(x_*).$$

Then $\sum_{i\in I(x_*)} \lambda_i^* = 1$ and $\lambda_i^* \geq 0$ for all $i \in I(x_*)$. Take any $i \in I(x_*)$, because the target sets are disjoint assumed by the standing assumptions, $d(x_*;\Omega_i) > 0$ and we have

$$x_i^* := \lim_{k\to\infty} \frac{x_k - \mathcal{P}(x_k;\Omega_i)}{g_i(x_k,p_k)} = \frac{x_* - \mathcal{P}(x_*;\Omega_i)}{d(x_*;\Omega_i)} = \partial d(x_*;\Omega_i).$$

Thus,

$$0 = \lim_{k\to\infty} \nabla\mathcal{D}(x_k,p_k) = \sum_{i\in I(x_*)} \lambda_i^* x_i^*.$$

This is equivalent to

$$0 \in \partial\mathcal{D}(x_*) = \text{co} \cup \left\{\partial d(x_*;\Omega_i) \mid i \in I(x_*)\right\},$$

which follows from a well-known subdifferential formula for "max" function; see [9]. Therefore, $x_*$ is an optimal solution of problem (1). The proof is complete. □

# 5 Implementation with the Nesterov Smooth Algorithm

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,1}$ function such that

$$|\nabla f(x) - \nabla f(y)| \leq \ell\|x - y\| \text{ for all } x, y \in \mathbb{R}^n.$$

For $x \in \mathbb{R}^n$, let

$$T(x) := \arg\min{}_y \langle\nabla f(x), y - x\rangle + \frac{\ell}{2}\|x - y\|^2.$$

Let $d(x)$ be a strongly convex function such that $d(x_0) = 0$ and

$$d(x) \geq \frac{\sigma}{2}\|x - x_0\|^2 \text{ for all } x \in \mathbb{R}^n.$$

To minimize the function $f$ on $\mathbb{R}$, the first order optimization scheme below was introduced by Nesterov in [7].

**Nesterov Smooth Algorithm**.

INPUT: $f$, $\ell$.
INITIALIZE: Choose $x_0 \in \mathbb{R}^n$.
For $k \geq 0$ do
  Compute $f(x_k)$ and $\nabla f(x_k)$.
  Find $y_k = T(x_k)$.
  Find $z_k = \arg\min_x \{\frac{\ell}{\sigma} d(x) + \sum_{i=0}^{k} \frac{i+1}{2}[f(x_i) + \langle \nabla f(x_i), x - x_i \rangle]\}$.
  Set $x_{k+1} = \frac{2}{k+3} z_k + \frac{k+1}{k+3} y_k$.

When adapting the Nesterov algorithm to the smallest intersecting ball problem, for simplicity, we choose $d(x) = \|x\|^2$, so $\sigma = 2$ and $x_0 = 0$. It is clear that

$$T(x_k) = x_k - \frac{\nabla f(x_k)}{\ell},$$

where $\ell = 2/p$. Moreover,

$$z_k = -\frac{1}{\ell} \sum_{i=0}^{k} \frac{i+1}{2} \nabla f(x_i).$$

Let us present a pseudo code for solving the smallest intersecting ball problem.

**MM Smoothing Method with Nesterov Algorithm.**

INPUT: $m$ target sets $\Omega_i$, $i = 1, \ldots, m$.
INITIALIZE: $\sigma \in (0, 1)$, $y_0 \in \mathbb{R}^n$ and $p_0 > 0$.
Set $k = 0$.
**repeat**
  Using the Nesterov iterations to solve the following with a stopping criterion
  $$\min_{x \in \mathbb{R}^n} \{G(x|y_k|p_k)\},$$
  and denote its solution by $x_k$.
  Set $p_{k+1} = \sigma p$, $y_{k+1} = x_k$.
  Set k:=k+1.
**until a stopping criterion is satisfied**.

# References


[1] D. Cheng, X. Hu, and C. Martin, On the smallest enclosing balls, Commun. Inf. Syst. **6** (2006), 137–160.

[2] E. Chi and K. Lange, A look at the generalized Heron problem through the lens of majorization-minimization, Amer. Math. Monthly, to appear.

[3] L. Drager, J. Lee, C. Martin, On the geometry of the smallest circle enclosing a finite set of points, J. Franklin Inst. **344** (2007), 929–940.

[4] B.S. Mordukhovich, N.M. Nam, J. Salinas, Solving a generalized Heron problem by means of convex analysis, Amer. Math. Monthly **119** (2012), 87–99.

[5] N.M. Nam and N. Hoang, A generalized Sylvester problem and a generalized Fermat-Torricelli problem, to appear in Journal of Convex Analysis.

[6] N.M. Nam, N. T. An, J. Salinas, Applications of convex analysis to the smallest intersecting ball problem, J. Convex Anal., J. Convex Anal. **19** (2012), 497–518.

[7] Yu. Nesterov, Smooth minimization of non-smooth functions, Math. Program. **103** (2005), 127–152.

[8] F. Nielsen, R. Nock, Approximating smallest enclosing balls with applications to machine learning. Internat. J. Comput. Geom. Appl. **19** (2009), 389–414.

[9] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[10] J.J. Sylvester, A question in the geometry of situation. Quarterly Journal of Pure and Applied Mathematics 1:79 (1857).

[11] E. Welzl, Smallest enclosing disks (balls ellipsoids). H. Maurer, editor, Lecture Notes in Comput. Sci. **555** (1991), 359–370.

[12] G. Zhou, K. C. Toh, and B. Sun, Efficient algorithms for the smallest enclosing ball problem, Comput. Optim. Appl. **30** (2005), 147–160.